Assessing Critical Attributes of Prospective Teachers: Implications for Selection into Initial Teacher Education Programs

Robert M. Klassen and Lisa E. Kim

University of York

Correspondence concerning this article should be addressed to Robert M. Klassen, Department of Education, University of York, Heslington, York, United Kingdom YO10 5DD

Email: robert.klassen@york.ac.uk

Author Note

This research was supported by a Consolidator Grant from the European Research Council awarded to the first author.

Assessing Critical Attributes of Prospective Teachers: Implications for Selection into Initial Teacher Education Programs

Abstract

In this article, we present an overview of a research program that is focused on reliably and validly measuring non-cognitive attributes associated with teacher effectiveness for the purpose of selection into ITE programs. The novel contribution of this program of research is that it builds on research and theory from educational psychology, methodology from organizational psychology, and findings from selection practices in medical education to address a critical educational problem. This article contributes four new insights into the selection of teachers and prospective teachers. First, we provide an overview of a dynamic interactionist view of the formation of effective teachers. Second, we describe how context, learning opportunities, and personal characteristics work together to influence student and teacher outcomes. Third, we introduce a conceptual model of how teacher selection measures are related empirically and conceptually to teacher effectiveness. Finally, we show how theory and research on teacher selection might be implemented in a six-stage selection process. Research into teacher selection has the potential to contribute to our understanding of the psychological factors associated with teacher effectiveness and to improve the quality of teachers entering the profession.

Assessing Critical Attributes of Prospective Teachers: Implications for Selection into Initial Teacher Education Programs

Teachers play a more important role in influencing student learning than almost any other factor. In Hattie's (2009) review of the factors associated with student learning and achievement, 'teacher factors' made a stronger contribution to student learning than home, curriculum, student, or school factors. Although teachers are linked with improvements in student learning, they are not "interchangeable parts" (Weisberg et al., 2009, p. 9), and individual differences in teacher effectiveness mean that student outcomes are significantly and reliably associated with *who* is doing the teaching. Atteberry's work on effectiveness within large cohorts of newly-hired teachers (Atteberry, Loeb, & Wyckoff, 2015) shows that teachers' *relative* effectiveness is stable; that is, effectiveness measured at the very start of a career is predictive of effectiveness later in a career, and these measures are especially predictive for those who initially display the highest and lowest levels of effectiveness. The evidence for individual differences in teacher effectiveness is persuasive (Atteberry et al., 2015; Chetty, Friedman, & Rockoff, 2014; Hanushek & Rivkin, 2012; Xu, Özek, & Hansen, 2015), yet school systems are often reluctant to publicly acknowledge that individual teachers vary in their levels of effectiveness (Scott & Dinham, 2008; Weisberg et al., 2009).

Identifying the cognitive and non-cognitive attributes[1] associated with teacher effectiveness is a question that educational psychology researchers have tackled for several decades. In the UK, the Department for Education guidance for initial teacher education (ITE) programs mandates that attention is paid to cognitive ("appropriate intellectual and academic capabilities") *and* non-cognitive ("personal qualities, attitudes, ethics and values") attributes (Department for Education, 2016, p. 11) attributes of candidates. In other settings, including

[1] We use the term non-cognitive attributes to refer to within-person variables including traits, motivation, personality, beliefs, attitudes, and dispositions. The term 'non-cognitive' is used in contrast to 'cognitive' factors such as subject area knowledge or reasoning ability that are routinely collected (academic transcripts, SAT/GRE) and used to inform selection decisions. In some literature, the term 'non-academic' is used in place of the term 'non-cognitive.'

high-performing education systems such as Finland and Singapore, considerable attention is paid to measuring cognitive and non-cognitive attributes for the purpose of selection into ITE programs (Sahlberg, 2015; Sclafani, 2015). For educational psychologists—most comfortable working with theory and learning processes (Berliner, 2006)—applying research findings to real world problems, such as the selection of candidates for ITE programs, presents a real challenge.

Educational psychologists have found non-cognitive attributes such as self-efficacy (Klassen & Tze, 2014), personality (Kim & MacCann, 2016; Robertson-Kraft & Duckworth, 2014), and teachers' beliefs about the nature of teaching and learning (Fives & Buehl, 2008; Sosu & Gray, 2012) to be associated with teacher and student outcomes. But can attributes measured at selection into ITE reliably predict a teacher's future effectiveness? Cognitive attributes (e.g., academic achievement, literacy and numeracy skills, subject area knowledge, and pedagogical knowledge) are widely and systematically assessed in high-stakes selection settings through, for example, examination of school records or administration of standardized tests. Non-cognitive attributes, on the other hand, are not as widely and systematically assessed and are difficult to measure given assessment problems, such as response biases and faking (Johnson & Saboe, 2011). The selection of prospective teachers benefits from multiple predictors because teacher effectiveness is multidimensional and may not be easily predicted using a single predictor (Harris & Rutledge, 2010; Hattrup, 2012). The purpose of this article is to present a research program focused on the practical problem of the selection of prospective teachers. The research program is built on theory and research not just from educational psychology, but also from organizational psychology and from well-developed selection practices in other professional disciplines.

Selection Research in Other Disciplines

Selection research is well established in fields outside of education. Ryan and Ployhart's (2014) review of the last 100 years of selection research noted that selection research is a mature field that continues to answer fundamental questions (*What should be*

assessed? How should we assess it?) with three key trends. First, more research is focusing on expanding the criterion space, that is, establishing a broader definition of what 'success' at work or in training means and how it can be conceptualized and measured. Second, researchers are continuing to develop approaches to measure diverse non-cognitive constructs, including emotional intelligence, social skills, integrity, and personality. Third, methodological advances in selection research include a new emphasis on situational judgment tests (SJTs) as a viable methodology to predict important outcomes in a range of contexts. Ryan and Ployhart also forecast future developments in selection research to include a shift from a Western-centric to a multicultural view of the key attributes targeted in selection, and an increasing emphasis on technological innovation in selection methods.

Selection research in medicine. Research on selection into medical training programs benefits from a well-established research base. Longitudinal predictive validity studies show that non-cognitive attributes assessed at the point of entry into a training program are significantly associated with academic and professional effectiveness several years after selection (e.g., Lievens & Sackett, 2012). Measures of non-cognitive attributes tend to provide *incremental* validity—i.e., a significant increase in prediction—over and above cognitive predictors such as achievement test scores and educational background (e.g., Patterson, Ashworth, Kerrin, & O'Neill, 2013). Although the selection landscape in medicine differs in important ways from that in education (e.g., level of competition among candidates, work context, and cost of training), research on selection practices for medical training can provide useful insights when designing research on selection practices for ITE programs.

A recent review systematically evaluated the methods used to measure cognitive and non-cognitive attributes of candidates for selection into medical training (Patterson et al., 2016). Of the eight selection methods identified to assess cognitive and non-cognitive attributes (aptitude tests, academic records, personal statements, reference letters, SJTs, personality and emotional intelligence assessments, traditional interviews, multiple mini-interviews, and selection centres), four methods were deemed to be the most effective and fair

methods. For measurement of cognitive attributes, examination of academic records and aptitude tests (e.g., MCAT or UKCAT) was found to be effective and fair. For measurement of non-cognitive attributes, SJTs, multiple mini-interviews, and selection centres were found to be effective (higher predictive validities) and fair (leas prone to bias). Less effective selection methods included reference letters, personal statements, and traditional interviews, with sparse research either for or against the use of personality assessments and emotional intelligence for selection. The review concluded that although considerable progress had been made in selection research during the period covered (1997-2015), a distinct lack of theory-informed research was inhibiting the development of a richer understanding of how cognitive and non-cognitive attributes contribute to competence and career success. We suggest that ITE selection research has been inhibited by a similar lack of theory-informed research.

Are 'Good Teachers' Born or Made?

The aim of ITE selection is to identify candidates with characteristics (whether fixed or malleable) that are associated with effective teaching. When ITE selectors are choosing a restricted number of candidates from among a pool of candidates, they use two evaluative processes: (a) they evaluate candidates' background factors, cognitive attributes, and non-cognitive attributes, and (b) they evaluate candidates' potential for developing these attributes during the ITE program and early teaching career. Some attributes, such as self-efficacy (Gutman & Schoon, 2013), may be more malleable than other attributes, such as personality (Tucker-Drob & Briley, 2014). Selection involves determining the extent to which candidates display the desired attributes (the 'born teacher') and/or the potential to develop the desired attributes (the 'made teacher'). Educational researchers have labelled the belief that teachers are born and not made a "damaging myth" that results in policies that rely on "some kind of prenatal alchemy" (Darling-Hammond, 2006, p. ix) to identify and prepare effective teachers. Scott and Dinham (2008) labelled this belief a "nativist myth" (p. 115) that is widely—and, in their view, incorrectly—held by many teachers. However, many educational psychologists (e.g., Kunter, Kleickmann, Klusmann, & Richter, 2013) believe that

personal characteristics such as intelligence, motivation (including competence beliefs), and personality interact with environmental factors such as professional learning opportunities to shape pre-service and practicing teacher professional outcomes.

Three views of the formation of effective teachers. The born-or-made question has important implications for teacher selection. Borrowing from the language of Dweck's implicit theories of intelligence (1986), some hold an *incremental view* about the factors leading to teacher effectiveness while others hold an *entity view* of the attributes. In Figure 1, we highlight the relevance of selection from three viewpoints: (a) the incremental view (referred to as the *qualification hypothesis* by Kunter et al., 2013) which reflects the 'good teachers are made' argument; (b) the entity view (referred to as the *individual aptitude hypothesis* by Kunter et al.), which reflects the 'good teachers are born' argument; and (c) the *dynamic interactionist* view, which recognizes that good teachers develop through the interaction of relatively stable personal characteristics with environmental factors. For those holding an incremental view, selection is not very important because they believe that key teacher attributes and skills can be developed through effective teacher training and professional development. For those holding an entity view, selection is *everything* because personal characteristics are largely immutable and hence play a key role in developing an effective teaching force. In this view, teacher training and professional development are less important than selecting teachers with the right set of personal characteristics.

Figure 1 here

In the dynamic interactionist view, individual differences in teachers' personal characteristics interact with environmental and learning factors to influence effectiveness in the classroom. Levels of teacher effectiveness increase over time as teachers gain experience (Atteberry et al., 2015) and as they learn new approaches to teaching and interacting with students. The individual differences in teacher effectiveness are influenced by within-person cognitive factors (e.g., verbal ability, numerical ability), non-cognitive factors (e.g., personality, motivation, and beliefs about teaching and learning), and by the quality and

quantity of learning opportunities provided. We believe that the dynamic interactionist view acknowledges the importance of the cognitive and non-cognitive attributes that candidates bring into ITE programs and into the profession, but also the interactions of these important personal characteristics with environmental factors. Thus, for those holding an interactionist view, selection is important because within-person characteristics interact with high quality training and development opportunities.

The Development of Teacher Effectiveness

Teachers differ in many ways, some of which are fleeting (e.g., moods), and others which are more entrenched (e.g., reasoning ability and personality). Although cognitive and non-cognitive attributes are influenced by environmental factors (e.g., learning opportunities and socialization), there are also individual differences with a biological basis, with high levels of heritability for intelligence, personality, and motivation (Gottschling, Spengler, Spinath, & Spinath, 2012; Krapohl et al., 2014). Cognitive factors such as intelligence account for a large part of the heritability of academic achievement, but non-cognitive factors are also heritable. Genetic influences account for more than half of the correlation between academic achievement and non-cognitive factors such as self-efficacy and personality (Krapohl et al., 2014). In the workplace, cognitive abilities have consistently been shown to predict measures of job performance, especially for work roles that are complex and require active information processing and managing simultaneous mental tasks (Murphy, 2012). Research on teacher effectiveness has shown the importance of non-cognitive attributes in the prediction of teacher effectiveness (e.g., Klassen & Tze, 2014), but how these attributes develop and influence teacher effectiveness is less well known.

Making a decision about admissions for teacher training programs represents a prediction about future effectiveness. In his book *Talks to Teachers,* William James (1899) spoke of the *art* of teaching, and proposed that good teachers display "an additional endowment altogether, a happy tact and ingenuity to tell us what definite things to say and do when the pupil is before us." ITE admissions teams' selection decisions may be guided by

government policies (e.g., Department for Education, 2016). Nonetheless, their primary goal is to identify candidates who display the personal attributes *and* the subject and pedagogical knowledge that are believed to lead to successful outcomes. Teacher effectiveness emphasizes the individual teacher as a causal factor in student learning whereas teaching effectiveness refers to the practices associated with successful outcomes, which may be acquired through training and professional development.

Research on teacher effectiveness conducted by Kunter and colleagues in their COACTIV program of research (Kunter et al., 2013) is built on a dynamic interactionist view of teacher effectiveness. In contrast, other models of teacher competence and teacher effectiveness exclusively highlight the role of learning opportunities—teacher education and professional learning—in the development of effectiveness, with little attention paid to within-teacher factors (e.g., Muijs et al., 2014). The COACTIV model of teachers' professional competence described by Kunter et al. proposes that competence develops over time through the provision of learning opportunities, but that competence is also influenced by critical personal characteristics that are present at entry into teacher training and practice. In this model, education systems and specific school context influence all aspects of teaching and learning through their relationship with learning opportunities, professional competence, and professional practice. In turn, these factors influence student and teacher outcomes. The contextual factors in this model, such as the macro-level educational system and specific institutional characteristics, provide learning opportunities that interact with existing characteristics of teachers. In terms of the born-or-made question, the answer, perhaps not surprisingly, is "both." The active engagement in and reflections on learning opportunities are not just dependent on the qualities of the opportunity, but are also influenced by the characteristics of the individual to whom the opportunity is presented.

In Figure 2, we present an adaptation of the COACTIV model of the development of teacher effectiveness for prospective teachers. In an ITE context, contextual factors (e.g., overarching national education system, specific characteristics of the ITE program, and

placement school characteristics) play a role in influencing the available learning opportunities that influence teaching competence and teaching behaviors. At the same time, personal characteristics also play a role in shaping teaching competence and teaching behaviors that lead to student and teacher outcomes. The selection process into an ITE program provides an opportunity to consider personal characteristics that will interact with contextual factors that together influence teacher effectiveness. Once candidates are admitted into an ITE program, the focus changes from identifying personal characteristics to providing appropriate learning opportunities. The personal characteristics (cognitive and non-cognitive attributes) of pre-service teachers provide a foundation that does not just influence the development of professional competence and professional practice, but also influences how they engage in available learning opportunities. Kunter et al.'s model recognizes both entity and incremental views of teacher effectiveness and provides a theoretical foundation to explain individual differences in teacher effectiveness.

Figure 2 here

Non-Cognitive Attributes Associated with Teacher Effectiveness

Research in education and psychology shows that multiple factors are related to teacher effectiveness. These factors include: (a) background factors such as educational record; (b) cognitive attributes such as subject knowledge and expertise, literacy and numeracy skills, pedagogical knowledge, and reasoning abilities; and (c) non-cognitive attributes such as self-efficacy, personality, and beliefs about knowledge. Kunter et al. (2013) showed that non-cognitive attributes (motivation and beliefs about teaching and learning) make an incremental contribution to successful teaching beyond pedagogical content knowledge. The factors related to teacher effectiveness are multifaceted, with non-cognitive attributes making a contribution over-and-above background and cognitive factors. In this section, we review three key non-cognitive attributes that have been shown to be associated with teacher effectiveness.

Self-efficacy. Motivation is defined as a contextualized and responsive set of wishes, desires, or underlying beliefs that influence people's movement towards attainment of goals (Pintrich & Schunk, 2002). Self-efficacy is a motivation construct—sometimes referred to as a *competence belief*—that concerns people's beliefs about their capabilities to carry out the courses of action required to accomplish particular goals. Motivation beliefs such as self-efficacy show some stability once established (Bandura, 1999), but may fluctuate during ITE (Klassen & Durksen, 2014). Overall, teachers who have higher levels of self-efficacy are more likely to be rated by classroom observers as being instructionally effective (Klassen & Tze, 2014).

Teachers' self-efficacy—a teacher's belief in the capability to influence student outcomes—has received considerable research attention, with a growing acknowledgement of its influence on student and teacher behaviors. An individual's self-efficacy beliefs operate as a motivation variable by increasing effort and persistence of the behaviors required for successful goal completion. Research shows that teachers' self-efficacy is related to job satisfaction (Klassen & Chiu, 2010), level of stress (e.g., Klassen & Chiu, 2011), and quality of relationships with students (Rimm-Kaufman & Hamre, 2010). Klassen and Tze (2014) conducted a meta-analysis of 43 studies representing 9,216 participants, investigating the link between teachers' psychological characteristics and externally measured teaching effectiveness. The relationship between teachers' self-efficacy and observed teaching performance was significant and of medium magnitude (r = .28, equivalent to Cohen's d of .58). Growing evidence on the association between teacher efficacy and teaching and learning outcomes is highlighting some new areas of research focus.

Although motivation may show some variation over time, the underlying patterns of motivation may be stable. Watt and Richardson (2008) measured the motivation of pre-service teachers during their teacher training programs in Australia. Using cluster analysis, the researchers found that a sizable proportion of participants with low motivation, the so-called 'lower engaged desisters,' showed little change in motivation profiles over the course of the

teacher training program. In other words, the group of pre-service teachers with low motivation were disaffected with teaching at the very beginning of their training and showed little change as their ITE program progressed, maintaining this disaffection during the first years of their careers. Pre-service teachers' motivation profiles over the course of an ITE program in the United States showed similar findings (Watt, Richardson, & Wilkins, 2014). The implications for ITE program candidate selection are clear: predictable and stable motivation profiles of pre-service and practicing teachers suggest that selection into teacher training and practice has long-term consequences.

Personality. Personality refers to non-cognitive attributes that tend to be expressed in the same way across situations and over time (McCrae & Costa, 1987). The 'Big Five' is the dominant personality framework, which posits that there are five key traits: conscientiousness, agreeableness, neuroticism, extraversion, and openness (John, Naumann, & Soto, 2008). Modern interactionist approaches suggest that how traits are expressed is shaped by the interaction between the person and the specific situation (Steyer, Schmitt, & Eid, 1999). Traits are not expressed in an invariant way: the expression of one's state is underpinned by an underlying latent trait that may be expressed differently according to the features of a particular context. Longitudinal research on personality suggests that traits identified in childhood and adolescence continue to show associations with behaviors and outcomes far into adulthood (Spengler et al., 2015). Research on teacher personality and effectiveness is not very well established, but Klassen & Tze's (2014) meta-analysis reported a modest but significant positive relationship between teachers' personality and objectively measured teaching effectiveness. Patrick (2011) found that students favored teachers who displayed higher levels of conscientiousness, openness, extraversion, and agreeableness (in descending order), but not neuroticism. Kim and MacCann (2016) showed that university students preferred courses taught by instructors with personality profiles closest to their self-described 'ideal instructor.' More research into how teacher personality is linked to teacher effectiveness is needed (Rimm-Kaufman & Hamre, 2010).

Teacher beliefs. The beliefs teachers hold about the nature of knowledge, learning, and ability (sometimes called epistemic or epistemological beliefs) are related to instructional behavior (Sosu & Gray, 2012). These beliefs—both implicit and explicit—shape teachers' classroom behaviors and the way that they interpret student behavior. Research on beliefs about teaching ability reveals important implications for teachers' development and resilience (Fives & Buehl, 2014). The beliefs of pre-service teachers act as filters when interpreting training content and experiences (Levin, 2015): including an assessment of teacher beliefs at the point of selection may be important for education programs in order to understand how candidates will respond to teacher training. Initial teacher education provides an opportunity to develop candidates' beliefs about teaching and learning (Buehl & Fives, 2009), but how effective ITE is in changing pre-service teachers' beliefs is not well established. The relationship between epistemic beliefs and teacher effectiveness is not firmly established, and to this point little is known about how these beliefs are amenable or resistant to change through training and professional practice.

Teacher Selection: Practice and Research

The purpose of selection for ITE programs is to identify the cognitive and non-cognitive attributes, and background factors believed to be critical for success in the program and in subsequent teaching. But many of the attributes cannot be measured directly (e.g., personality), and must be inferred from imperfect measures. Figure 3, adapted from Binning and Barrett's (1989) classic personnel selection model (and Ployhart and Schneider's [2012] re-interpretation of the classic model) shows how selection consists of a series of inferences based on theoretical and empirical relationships. In this model, latent constructs are represented by circles, observed variables by rectangles, and inferences are represented by dashed arrows.

Figure 3 here

Consider an ITE program that decides to measure the personality construct of *conscientiousness* in candidates in their selection process. Arrow 4 represents the empirical

relationship between a selection measure (e.g., measure of conscientiousness) and an outcome measure (e.g., rating on teaching practicum), and is usually assessed as the correlation between the selection measure (i.e., predictor) and scores on a teaching outcome measure (i.e., criterion outcome). Arrow 2 mirrors the relationship between the predictor and outcome but denotes its theoretical or latent relationship. Arrows 1 and 3 represent construct validity: arrow 1 represents the extent to which the selection measure (predictor) represents the construct of interest, and arrow 3 represents the construct validity of the outcome measure, and whether the outcome, usually some kind of measure of teacher effectiveness, represents the person's 'true' teaching effectiveness. Arrow 5 lies at the heart of the selection process and indicates the degree to which scores from an imperfect selection measure (e.g., personality test, letter of reference, face-to-face interview) are associated with 'true' differences in teacher effectiveness, imperfectly measured by an outcome measure (e.g., grade on teaching practicum). Although the relationship between the actual measure used in selection and the latent teacher effectiveness variable cannot be directly assessed, it can be inferred through the other relationships (arrows) described in the model. Research on teacher selection fails to capture the richness of the hypothesized relationships among variables if it only focuses on the correlation between predictor measure and outcome measure.

Challenges in high-stakes selection. When a non-cognitive attribute (e.g., construct of motivation or personality) is measured in a research project, the process is (relatively) straightforward: participants are asked to report their levels of motivation or personality, and the researcher assumes that the reported scores accurately reflect the targeted construct. Participants may not always respond honestly, but the 'cloak of anonymity' provides little incentive to distort their responses. However, assessing non-cognitive attributes in a high-stakes setting is much more difficult because candidates may be motivated to distort their responses in order to improve their chances of success in the selection exercise (Johnson & Saboe, 2011). Candidates in a selection process have strong motivation to provide responses that they believe will show them in the most positive light and that will increase their odds of

succeeding at gaining a place on a training program. *Differential* faking is particularly concerning, where some candidates might inflate their scores more than others, thus affecting the fairness of the selection process. For example, when all candidates engage in a 'fake-good' response pattern to the same degree on a personality test, everyone benefits to the same degree. The problem becomes more serious when faking patterns favor certain candidates or groups of candidates, for example, candidates who are able to afford coaching for the test, which alters the fairness of selection decisions (Ryan & Ployhart, 2014).

A word about effect sizes. Selection methods are never perfect predictors of work performance. In the complex and messy world of education, predicting outcomes (student outcomes or teacher effectiveness) is incredibly challenging, and effect sizes for predictor variables rarely reach the heights of those in more controlled research settings. In educational research, Hattie (2009) suggests the following effect size descriptions: $d = .20$ (roughly equivalent to $r = .10$) describes a small effect, $d = .40$ ($r = .20$) describes a medium effect, and $d = .60$ ($r = .30$) describes a large effect. Interpretations of the *practical value* of effect sizes in education are fluid: Coe (2002) proposed that an effect size of $d = .10$ (roughly $r = .05$) can have important educational implications if the effect can be applies to all students (i.e., as in an effect involving teachers) and is cumulative over time. Individual selection methods rarely display effect sizes above $d = .30$, but the use of multiple selection methods offers the possibility of significant added value (incremental validity gain), resulting in improved selection decisions.

Selection practices in education. Although some teacher educators may believe that teachers are made, not born, most ITE programs implicitly acknowledge the importance of existing attributes through the assessment of cognitive and non-cognitive attributes during the selection process. We recently surveyed 74 university-based ITE programs in England and Wales to understand how cognitive and non-cognitive attributes were assessed for selection (Klassen & Dolan, 2015). All of the programs assessed cognitive attributes through a review of academic records (A-levels, GCSE grades in English and Math, and university degree

class, i.e., 1st, 2:1, 2:2, etc.) and all of the programs assessed non-cognitive attributes in some way, using individual and group interviews (97%), assessment of social behaviors through group activities (62%), and formal personality tests (3%). In educationally high-performing systems such as Finland and Singapore, selection methods include cognitive and non-cognitive assessments. Selection into ITE programs in Finland includes evaluation of personality and interpersonal skills using a range of interviews and tests (Sahlberg, 2015), and into ITE programs in Singapore includes an evaluation non-academic attributes including motivation, passion, values, and commitment to teaching (Sclafani, 2015).

Research on selection into ITE programs suggests that selection methods tend to be static (fixed over time, with little change in the kinds of methods used), and lacking in evidence of effectiveness (e.g., Caskey, Peterson, & Temple, 2001). Assessment of cognitive and non-cognitive factors for ITE selection is common not just in the UK, but also internationally (e.g., Heinz, 2013; OECD: Organisation for Economic Cooperation and Development, 2005; Sahlberg, 2015). Unfortunately, most selection practices for ITE programs are built on folk beliefs and external drivers such as time constraints and recruitment challenges, with scant evidence for the reliability and validity of many selection methods (Caskey, Peterson, & Temple, 2001).

Implicit and explicit measurement of non-cognitive attributes. For ITE programs, assessing candidates' cognitive attributes is relatively straightforward: schools and universities routinely assess and record academic progress, and standardized tests (e.g., SAT) are available to measure reasoning abilities. In contrast, assessing non-cognitive attributes is more difficult. Although some aspects of motivation, personality, and beliefs operate 'on the surface' or explicitly, other aspects operate implicitly, and are separate from people's awareness and control (Schultheiss & Brunstein, 2010). Personnel selection researchers Motowidlo and Beier (2010) propose an *implicit trait policies* theoretical framework that suggests that we can gain insight into implicit non-cognitive traits by asking an individual to judge the effectiveness of responses to situations designed to elicit targeted characteristics.

When faced with a personality test item, individuals draw on their ready-to-hand explicit beliefs (e.g., *I am generally agreeable)*. But behaviors are also influenced by implicit factors, which may not be readily accessible by items that tap only surface attributes. In education, teaching behaviors are influenced by a combination of implicit and explicit attributes.

Using Situational Judgment Tests in Selection Research

One key challenge for selectors into professional training programs—in education just as in medicine, law, business, or nursing—is how to reliably, validly, and efficiently measure non-cognitive attributes. Increasingly, selectors are choosing to use SJTs to capture key non-cognitive attributes associated with success in training and professional practice (Campion, Ployhart, & MacKenzie, 2014). SJTs are a measurement method designed to assess candidates' judgments of the implications of behaving in certain ways in response to contextualized scenarios. For ITE programs, the contextualized scenarios typically take place in the classroom. After presentation of a scenario, candidates are asked what they should do in the situation, and then to choose responses from a set of response options. Although SJTs can be construed as a type of written structured interview, they offer the advantage of wider sampling of classroom situations, a scoring key that is standardized, and the capacity to screen large numbers of candidates in an economical and efficient manner. The administration format of SJTs can be varied, such as being presented on paper-and-pencil, computer, or video. The development of SJT content is typically based on job analysis and through gathering 'critical incidents' from those already in the job (Whetzel & McDaniel, 2009). Experienced professionals, or 'subject matter experts,' contribute to the development process by generating response options (Lievens & Sackett, 2012). Scoring keys, which reflect the effectiveness of the response options, are established through consensus with a panel of subject matter experts.

SJTs are designed to measure *implicit trait policies*; that is, the tendency individuals have to express traits in certain ways under particular contexts (Motowidlo & Beier, 2010). According to this theory—similarly conceptualized as tacit knowledge in

Sternberg's theory of successful intelligence (Elliott, Stemler, Sternberg, Grigorenko, & Hoffman, 2011)—those who are more experienced in a particular job are more likely to implicitly understand optimal behavioral responses. However, novices with limited experience also have partial knowledge about effective response patterns, based on their implicit traits and understanding of the kinds of behaviors that are likely to be most appropriate in SJT scenarios (Motowidlo & Beier, 2010). Candidates for ITE programs may have pre-existing beliefs about how to react to classroom challenges (e.g., how to manage challenging student behaviors), based on the knowledge gained from their own life experiences, even when they do not have direct experience with teaching. These existing beliefs, or implicit trait policies, may change as candidates gain pedagogical knowledge and teaching experience, but their 'trait policies' remain as influences of teaching behaviors.

In comparison to conventional non-cognitive assessment methodologies, SJTs display stronger face and content validity due to their close correspondence with the work-related situations that they describe (Whetzel & McDaniel, 2009). The interest in SJT methodologies is due to the promise of predictive validity: SJTs administered at admissions to medical school predicted job performance nine years later (r = .22; Lievens & Sackett, 2012). A meta-analysis comparing SJT validities reported that SJTs measuring interpersonal attributes showed a mean validity coefficient of .25, those measuring conscientiousness showed a mean coefficient of .24, and heterogeneous composite SJTs showed a mean coefficient of .28 (Christian, Edwards, & Bradley, 2010). A previous large-scale meta-analysis of SJT validity (N = 24,756) using mostly concurrent validity studies showed a validity coefficient of .26 (McDaniel, Hartman, & Whetzel, 2007).

There are some limitations on how SJTs might be implemented in ITE programs. SJTs are relatively inexpensive to administer, but there can be a high cost incurred in their development (Patterson et al., 2016). ITE program selectors may be resistant to include quantitative measurement of non-cognitive attributes, with the belief that teaching is a unique profession that does not benefit from methods used in other fields (Harris & Rutledge, 2010).

Finally, SJT development can lack a strong theory base, with difficulty integrating well-developed theories into SJT content (Campion et al., 2014). In spite of these limitations and the extant research on SJTs, the method provides a promising avenue for measurement of non-cognitive attributes for selection into ITE programs.

A Theory- and Research-Driven Multi-Stage Selection Process

Recognizing that selection for ITE programs can be limited by restrictive government policies, practical exigencies, and a lack of exposure to selection research, we propose a theory- and research-informed process for selection into ITE programs that is guided by theory from organizational psychology and research on selection other professional fields. Over the last four years, we have begun to develop teacher selection models based on selection research in other disciplines (e.g., Klassen et al., 2017; Klassen, Durksen, Rowett, & Patterson, 2014). In Figure 4 we provide an example of how a multi-stage research- and theory-driven selection process could be implemented for choosing candidates for an ITE program. In contrast to the *ad hoc* and *information-poor* selection systems described in previous literature (e.g., Goldhaber, Grout, & Huntington-Klein, 2014), the model we describe is grounded in research on cognitive and non-cognitive attributes and teacher effectiveness. Furthermore, in contrast to current static models of selection methods, our model is underpinned by an iterative evaluative cycle that has the capacity to continuously refine the selection process and assessment.

Figure 4 here

In Stage 1, researchers and program leaders conduct an analysis of critical attributes necessary for success in an ITE program and for future teaching success. In this stage, program staff may liaise with mentor teachers to consider the personal characteristics—cognitive and non-cognitive—that will be targeted and included in the selection process and subsequent training program. In Stage 2, candidates' eligibility to enter the program is checked to assess whether they possess the appropriate academic record required for the program. In Stage 3, candidates' levels of cognitive and non-cognitive attributes are

evaluated. We suggest that this screening phase might include assessments of literacy and numeracy skills, non-cognitive attributes measured through SJTs, and, any other cognitive ability assessments that the ITE program staff may wish to assess (e.g., specific subject knowledge). The purpose of Stage 3 is to identify, in a cost-effective way, the candidates that will undergo the more labor-intensive (and expensive) selection process in the next stage. In Stage 4, evidence-supported methods are used to select the most promising candidates from the pool screened in Stage 3. We propose three Stage 4 selection methods: (a) simulated teaching (selection tasks which replicate criterion tasks are predictive of that criterion; Lievens & De Soete, 2015), (b) structured individual interviews or multiple mini-interviews (Patterson et al., 2016), and (c) SJTs (if not administered in Stage 3). In Stage 5, candidates are selected based on the aggregated data gathered in Stage 4. Stage 6 consists of a feedback loop designed to link teaching outcomes (gathered during and after ITE program) with the selection components of previous stages. The feedback data gathered in Stage 6 can be used to refine the overarching selection criteria (Stage 1), the screening measures from Stage 3, and the intensive selection methods used in Stage 4. The outcome measures in Stage 6 include evidence-supported measures of teacher effectiveness, including student achievement gains data, teaching observation data, and ITE and career attrition data (e.g., Pianta & Hamre, 2009).

New Developments in ITE Selection

Over the last four years, we have been working with ITE programs and national and state-level ministries of education in the UK (England, Scotland, and Northern Ireland), Australia, Finland, Lithuania, and Oman to develop theory- and research-based selection tools. In particular, we have focused on developing SJTs to assess candidates' non-cognitive attributes for selection into primary and secondary ITE programs. We followed a nine-step development process consisting of an analysis of teachers' roles and practices, multiple focus groups with stakeholders for item development and test construction, and piloting the SJTs with primary and secondary ITE candidates. Item development was based on a critical

incident approach, with content reflecting three composite domains that emerged from analysis of our extensive consultation with teachers and teacher educators: (a) empathy and communication, (b) organization and planning, and (c) resilience and adaptability. Results from pilot testing shows high levels of reliability and strong evidence of concurrent validity with currently used non-cognitive assessment methods (Klassen et al., 2017). We are currently comparing and revising non-cognitive attributes and SJT content across cultural settings, with results showing similarities in key non-cognitive attributes, but with an additional focus on *professional ethics* in Oman (Al Hashmi & Klassen, 2016), and *cooperation and fostering of community* in Finland (Metsäpelto, Poikkeus, & Klassen, 2016). Our cross-national comparisons show that there appear to be universal attributes associated with teacher effectiveness, i.e., it is agreed in all countries that resilience and adaptability are critically important, but that particular socio-cultural norms and expectations create context-specific emphasis on additional non-cognitive attributes.

Conclusions

In this article we presented an overview of a research program that is focused on reliably and validly measuring non-cognitive attributes associated with teacher effectiveness for the purpose of selection into ITE programs. The novel contribution of this program of research is that it builds on research and theory from educational psychology, methodology from organizational psychology, and findings from selection research in medical education to address a critical educational problem.

ITE program selection undoubtedly faces a wide range of logistical, ethical, and measurement challenges. We propose that educational psychology researchers address these critically important, but often neglected, real world challenges by addressing these questions:

1. What are the critical non-cognitive attributes linked to teacher effectiveness? Only a modest amount of research attention has been paid to linking constructs from educational psychology to objectively measured teacher effectiveness (Klassen & Tze, 2014).

2. How can we successfully measure theory-driven constructs, including motivation constructs such as self-efficacy and competence beliefs, in high-stakes selection settings?
3. How do these key non-cognitive attributes develop and change over time? How stable are patterns of motivation, personality, and beliefs from the point of selection into later career stages of teaching practice?
4. Can we develop theory-informed innovative technological tools that increase the fidelity of our selection processes? One option is to explore simulated environments using video-based or virtual reality environments that offer higher fidelity with classroom environments than paper-and-pencil selection tools (Rockstuhl, Ang, Ng, Lievens, & Van Dyne, 2015).

Developing reliable, valid, and fair assessment tools that adequately capture non-cognitive attributes for selection into ITE programs provides a real challenge for educational psychology researchers, but it is a challenge worth addressing. SJTs have been used successfully for selecting individuals into training programs in other professional disciplines. We propose that this methodology provides a promising way to measure non-cognitive attributes that may be used for selection into ITE programs. Although the research presented in this article is in its early stages, the preliminary results are promising and warrant further development and testing, with longitudinal and cross-cultural data particularly needed to show predictive relationships with teaching effectiveness. It is anticipated that with increasing international demand for improving educational systems, there will be a parallel demand for educational psychology researchers to apply their research to the challenging but critical task of selecting prospective teachers.

References

Al Hashmi, W., & Klassen, R. M. (2016, June). *To what extent can situational judgment tests measure the non-cognitive attributes of prospective teachers in Oman?* Paper presented at the bi-annual meeting of the European Association for Research on Learning and Instruction SIG 11 (Teaching and Teacher Education), Zurich, Switzerland.

Atteberry, A., Loeb, S., & Wyckoff, J. (2015). Do First Impressions Matter? Predicting Early Career Teacher Effectiveness. *AERA Open*, *1*(4), 1–23.

Bandura, A. (1999). Social cognitive theory of personality. *Handbook of Personality: Theory and Research*, 154–196.

Berliner, D. C. (2006). Educational psychology: Searching for essence throughout a century of influence. In P. A. Alexander & P.H. Winne (Eds.), *Handbook of educational psychology* (2nd ed., pp. 3-27). Mahwah, NJ: Lawrence Erlbaum.

Binning, J. F., & Barrett, G. V. (1989). Validity of personnel decisions: A conceptual analysis of the inferential and evidential bases. *Journal of Applied Psychology*, *74*, 478–494.

Buehl, M. M., & Fives, H. (2009). Exploring teachers' beliefs about teaching knowledge: Where does it come from? Does it change? *Journal of Experimental Education*, *77*, 367–408.

Campion, M. C., Ployhart, R. E., & MacKenzie, W. I. (2014). The state of research on situational judgment tests: A content analysis and directions for future research. *Human Performance*, *27*, 283–310.

Caskey, M. M., Peterson, K. D., & Temple, J. B. (2001). Complex admission selection procedures for a graduate preservice teacher education program. *Teacher Education Quarterly*, *28*(4), 7–21.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *American Economic Review*, *104*, 2593–2632.

Christian, M. S., Edwards, B. D., & Bradley, J. C. (2010). Situational judgment tests:

Constructs assessed and a meta-analysis of their criterion-related validities. *Personnel Psychology*, *63*, 83–117.

Coe, R. (2002, September). *It's the effect size, stupid: What effect size is and why it is important.* Paper presented at the Annual Conference of the British Educational Research Association, Exeter, England.

Darling-Hammond, L. (2006). *Powerful Teacher Education: Lessons from Exemplary Programs*. San Francisco: Jossey-Bass.

Department for Education (2016). *Initial teacher training criteria and supporting advice.* Retrieved from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/508303/initial-teacher-training-criteria-supporting-advice.pdf

Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist*, *41*, 1040–1048.

Elliott, J. G., Stemler, S. E., Sternberg, R. J., Grigorenko, E. L., & Hoffman, N. (2011). The socially skilled teacher and the development of tacit knowledge. *British Educational Research Journal*, *37*, 83–103.

Fives, H., & Buehl, M. M. (2008). What do teachers believe? Developing a framework for examining beliefs about teachers' knowledge and ability. *Contemporary Educational Psychology*, *33*, 134–176.

Fives, H., & Buehl, M. M. (2014). Exploring differences in practicing teachers' valuing of pedagogical knowledge based on teaching ability beliefs. *Journal of Teacher Education*, *65*, 435–448.

Goldhaber, D., Grout, C., & Huntington-Klein, N. (2014). *Screen twice, cut once: Assessing the predictive validity of teacher selection tools. (CEDR Working Paper No. 9)*. Seattle, WA: Center for Education Data and Research. Retrieved from http://www.cedr.us/papers/working/CEDR%20WP%202014-9.1.pdf

Gottschling, J., Spengler, M., Spinath, B., & Spinath, F. M. (2012). The prediction of school

achievement from a behavior genetic perspective: Results from the German twin study on Cognitive Ability, Self-Reported Motivation, and School Achievement (CoSMoS). *Personality and Individual Differences*, *53*, 381–386.

Gutman, L. M., & Schoon, I. (2013). The impact of non-cognitive skills on outcomes for young people. *Education Endowment Foundation.* Retrieved from*: https://educationendowmentfoundation.org.uk/public/files/Publications/EEF_Lit_Review_Non-CognitiveSkills.pdf*

Hanushek, E. A., & Rivkin, S. G. (2012). The distribution of teacher quality and implications for policy. *Annual Review of Economics*, *4*, 131–157.

Harris, D. N., & Rutledge, S. A. (2010). Models and predictors of teacher effectiveness: A comparison of research about teaching and other occupations. *Teachers College Record, 112,* 914-960.

Hattrup, K. (2012). Using composite predictors in personnel selection. In N. Schmitt (Ed.) *The Oxford handbook of personnel assessment and selection* (pp. 297-319). Oxford: Oxford University Press.

Hattie, J. (2009). *Visible learning: A synthesis of 800+ meta-analyses on achievement*. Abingdon: Routledge.

Heinz, M. (2013). Tomorrow's teachers—selecting the best: An exploration of the quality rationale behind academic and experiential selection criteria for initial teacher education programmes. *Educational Assessment, Evaluation and Accountability*, *25*, 93–114.

James, W. (1899). *Talks to teachers on psychology: And to students on some of life's ideals*. Henry Holt and Company.

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative big five trait taxonomy. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of Personality: Theory and Research* (3rd ed., pp. 114–158). New York, NY: Guilford Press.

Johnson, R. E., & Saboe, K. N. (2011). Measuring implicit traits in organizational research: Development of an indirect measure of employee implicit self-concept. *Organizational Research Methods*, *14*, 530–547.

Kim, L. E., & MacCann, C. (2016). What is students' ideal university instructor personality? An investigation of absolute and relative personality preferences. *Personality and Individual Differences*, *102*, 190–203.

Klassen, R. M., & Chiu, M. M. (2010). Effects on teachers' self-efficacy and job satisfaction: Teacher gender, years of experience, and job stress. *Journal of Educational Psychology*, *102*(3), 741.

Klassen, R. M., & Chiu, M. M. (2011). The occupational commitment and intention to quit of practicing and pre-service teachers: Influence of self-efficacy, job stress, and teaching context. *Contemporary Educational Psychology*, *36*, 114–129.

Klassen, R. M., & Dolan, R. (2015, September). *Selection for teacher education in the UK and the Republic of Ireland: A proposal for innovation.* Presented at the meeting of the European Conference on Educational Research, Budapest, Hungary.

Klassen, R. M., & Durksen, T. L. (2014). Weekly self-efficacy and work stress during the teaching practicum: A mixed methods study. *Learning and Instruction*, *33*, 158–169.

Klassen, R. M., Durksen, T. L., Kim, L. E., Patterson, F., Rowett, E., Warwick, J., Warwick, P., & Wolpert, M. A. (2017). Developing a proof-of-concept selection test for entry into primary teacher education programs. *International Journal of Assessment Tools in Education*, 4, *96-114.*

Klassen, R.M., Durksen, T.L., Rowett, E., & Patterson, F. (2014). Applicant reactions to a situational judgment test used for selection into initial teacher training. *International Journal of Educational Psychology*, *3*, 104-125. doi: 10.4471/ijep.2014.07

Klassen, R. M., & Tze, V. M. C. (2014). Teachers' self-efficacy, personality, and teaching effectiveness: A meta-analysis. *Educational Research Review*, *12*, 59–76.

Krapohl, E., Rimfeld, K., Shakeshaft, N. G., Trzaskowski, M., McMillan, A., Pingault, J.-B.,

… Plomin, R. (2014). The high heritability of educational achievement reflects many genetically influenced traits, not just intelligence. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, 15273–15278.

Kunter, M., Kleickmann, T., Klusmann, U., & Richter, D. (2013). The development of teachers' professional competence. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Eds.), *Cognitive activation in the mathematics classroom and professional competence of teachers* (pp. 63–77). New York, NY: Springer.

Levin, B. B. (2015). The development of teachers' beliefs. In H. Fives & M. G. Gill (Eds.), *International handbook of research on teachers' beliefs* (pp. 48–65). New York, NY: Routledge.

Lievens, F., & De Soete, B. (2015). Situational judgment tests. In J. D. Wright (Ed.), *International Encyclopedia of the Social & Behavioral Sciences* (2nd ed., Vol. 22, pp. 13–19). Oxford: Elsevier.

Lievens, F., & Sackett, P. R. (2012). The validity of interpersonal skills assessment via situational judgment tests for predicting academic success and job performance. *The Journal of Applied Psychology*, *97*, 460–468.

McCrae, R. R., & Costa, P. T., Jr. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, *52*, 81–90.

McDaniel, M. A., Hartman, N. S., & Whetzel, D. L. (2007). Situational judgment tests, response instructions, and validity: A meta□analysis. *Personnel Psychology*, *60*, 63–91.

Metsäpelto, R.-L., Poikkeus, A.-M., & Klassen, R. M. (2016). *Developing student selection for initial teacher training: Adapting the situational judgment test in Finland*. Presented at the bi-annual meeting of the European Association for Research on Learning and Instruction SIG 11 (Teaching and Teacher Education), Zurich, Switzerland.

Motowidlo, S. J., & Beier, M. E. (2010). Differentiating specific job knowledge from implicit trait policies in procedural knowledge measured by a situational judgment test. *Journal*

of Applied Psychology, *95*, 321–333.

Muijs, D., Kyriakides, L., van der Werf, G., Creemers, B., Timperley, H., & Earl, L. (2014). State of the art-teacher effectiveness and professional learning. *School Effectiveness and School Improvement*, *25*, 231–256.

Murphy, K. R. (2012). Individual differences. In N. Schmitt (Ed.), *The Oxford Handbook of Personnel Assessment and Selection* (pp. 31–47). New York, NY: Oxford University Press.

OECD: Organisation for Economic Cooperation and Development. (2005). *Teachers matter: Attracting, developing and retaining effective teachers*. Paris, France: OECD Publications. Retrieved from http://www.oecd.org/education/school/attractingdevelopingandretainingeffectiveteachers-finalreportteachersmatter.htm

Patrick, C. L. (2011). Student evaluations of teaching: effects of the Big Five personality traits, grades and the validity hypothesis. *Assessment & Evaluation in Higher Education*, *36*, 239–249.

Patterson, F., Ashworth, V., Kerrin, M., & O'Neill, P. (2013). Situational judgement tests represent a measurement method and can be designed to minimise coaching effects. *Medical Education*, *47*, 220–221.

Patterson, F., Knight, A., Dowell, J., Nicholson, S., Cousans, F., & Cleland, J. (2016). How effective are selection methods in medical education? A systematic review. *Medical Education*, *50*, 36–60.

Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher* , *38*, 109–119.

Pintrich, P. R., & Schunk, D. H. (2002). *Motivation in Education: Theory, Research, and Applications*. Merrill.

Ployhart, R. E., & Schneider, B. (2012). The social and organizational context of personnel

selection. In N. Schmitt (Ed.), *The Oxford handbook of personnel assessment and selection* (pp. 48–67). New York, NY: Oxford University Press.

Rimm-Kaufman, S. E., & Hamre, B. K. (2010). The role of psychological and developmental science in efforts to improve teacher quality. *Teachers College Record*, *112*, 2988–3023.

Robertson-Kraft, C., & Duckworth, A. L. (2014). True grit: Trait-level perseverance and passion for long-term goals predicts effectiveness and retention among novice teachers. *Teachers College Record*, *116*(3), 1–27.

Rockstuhl, T., Ang, S., Ng, K.-Y., Lievens, F., & Van Dyne, L. (2015). Putting judging situations into situational judgment tests: Evidence from intercultural multimedia SJTs. *Journal of Applied Psychology*, *100*, 464–480.

Ryan, A. M., & Ployhart, R. E. (2014). A century of selection. *Annual Review of Psychology*, *65*, 693–717.

Sahlberg, P. (2015). *Finnish Lessons 2.0: What Can the World Learn from Educational Change in Finland?* (2nd ed.). New York, NY: Teachers College Press.

Schultheiss, O., & Brunstein, J. (2010). *Implicit Motives*. New York, NY: Oxford University Press.

Sclafani, S. K. (2015). Singapore chooses teachers carefully. *Phi Delta Kappan*, *97*(3), 8-13.

Scott, C., & Dinham, S. (2008). Born not made: the nativist myth and teachers' thinking. *Teacher Development*, *12*, 115–124.

Sosu, E. M., & Gray, D. S. (2012). Investigating change in epistemic beliefs: An evaluation of the impact of student teachers' beliefs on instructional preference and teaching competence. *International Journal of Educational Research*, *53*, 80–92.

Spengler, M., Brunner, M., Damian, R. I., Lüdtke, O., Martin, R., & Roberts, B. W. (2015). Student characteristics and behaviors at age 12 predict occupational success 40 years later over and above childhood IQ and parental socioeconomic status. *Developmental Psychology*, *51*, 1329–1340.

Steyer, R., Schmitt, M., & Eid, M. (1999). Latent state–trait theory and research in personality

and individual differences. *European Journal of Personality*, *13*, 389–408.

Tucker-Drob, E. M., & Briley, D. A. (2014). Genetic and environmental continuity in personality development: A meta-analysis. *Psychological Bulletin*, *140*, 949-979.

Watt, H. M. G., & Richardson, P. W. (2008). Motivations, perceptions, and aspirations concerning teaching as a career for different types of beginning teachers. *Learning and Instruction*, *18*, 408–428.

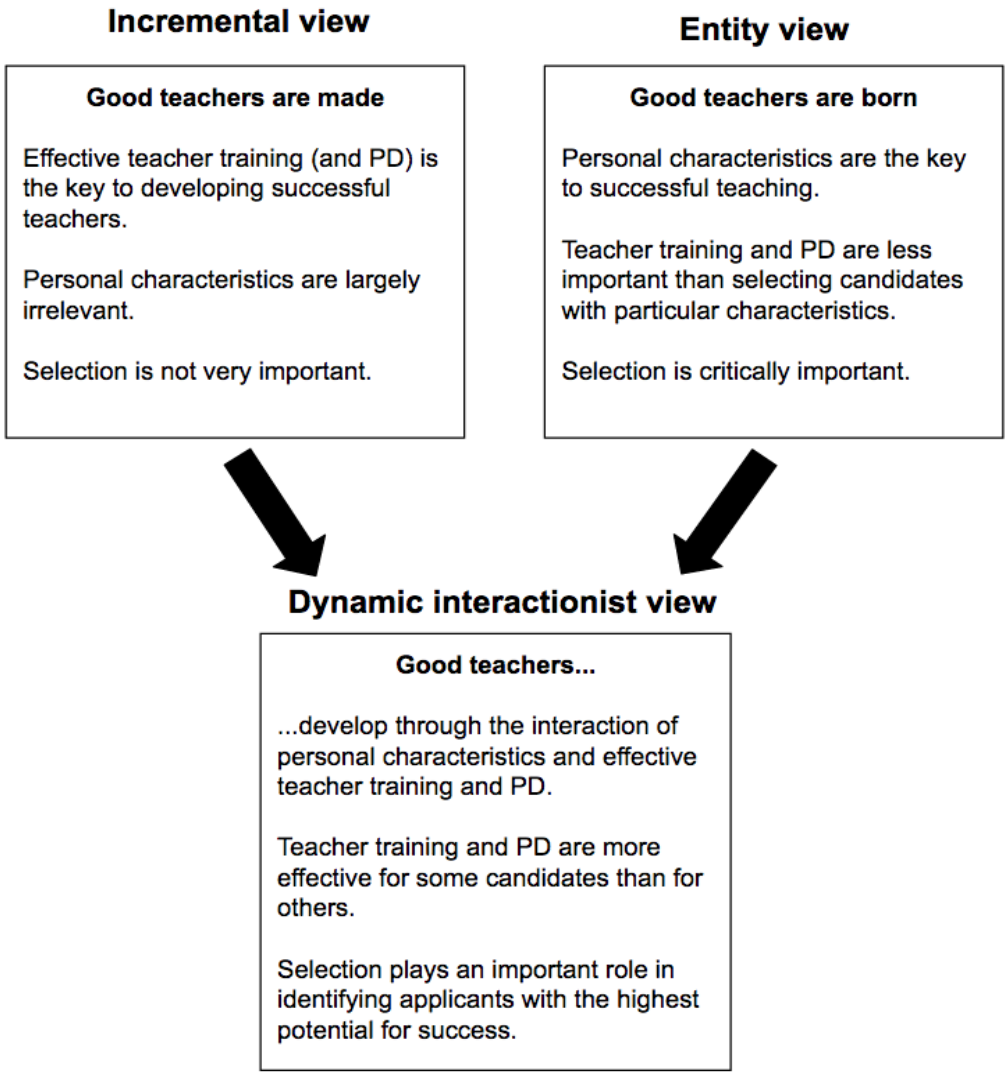
Watt, H. M. G., Richardson, P. W., & Wilkins, K. (2014). Profiles of professional engagement and career development aspirations among USA preservice teachers. *International Journal of Educational Research*, *65*, 23–40.

Weisberg, D., Sexton, S., Mulhern, J., Keeling, D., Schunck, J., Palcisco, A., & Morgan, K. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness*. The New Teacher Project. Retrieved from http://eric.ed.gov/?id=ED515656

Whetzel, D. L., & McDaniel, M. A. (2009). Situational judgment tests: An overview of current research. *Human Resource Management Review*, *19*, 188–202.

Xu, Z., Özek, U., & Hansen, M. (2015). Teacher performance trajectories in high-and lower-poverty schools. *Educational Evaluation and Policy Analysis*. Retrieved from http://epa.sagepub.com/content/37/4/458.short

Whetzel, D. L., & McDaniel, M. A. (2009). Situational judgment tests: An overview of current research. *Human Resource Management Review*, *19*, 188–202.



e 1. Three views on the development of effective teachers.

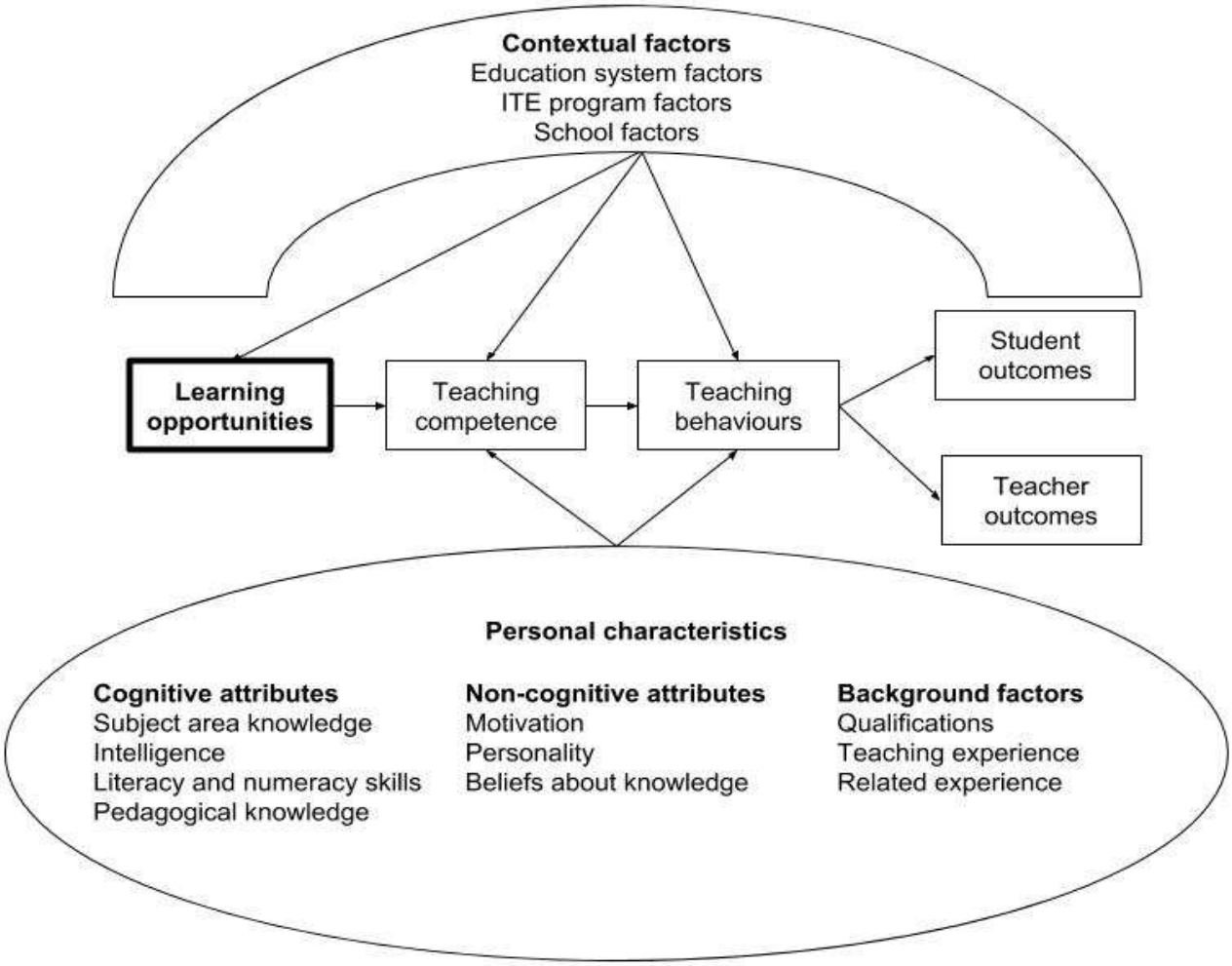


Figure 2. Model of the development of teacher effectiveness (adapted from Kunter et al., 2013).

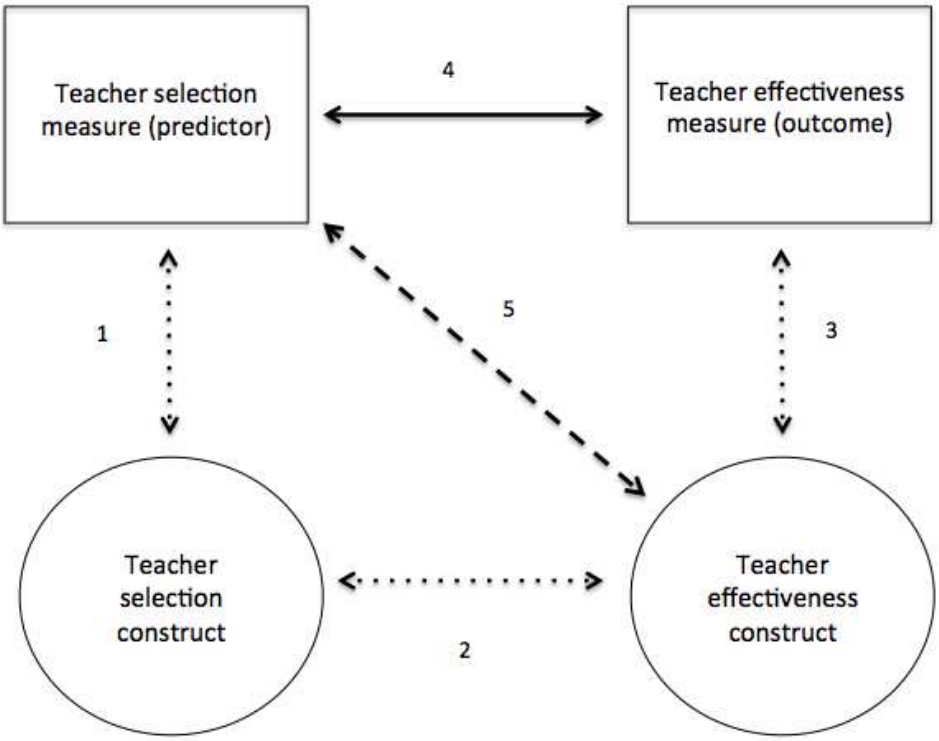


Figure 3. Teacher selection model (Adapted from Binning & Barrett, 1989).

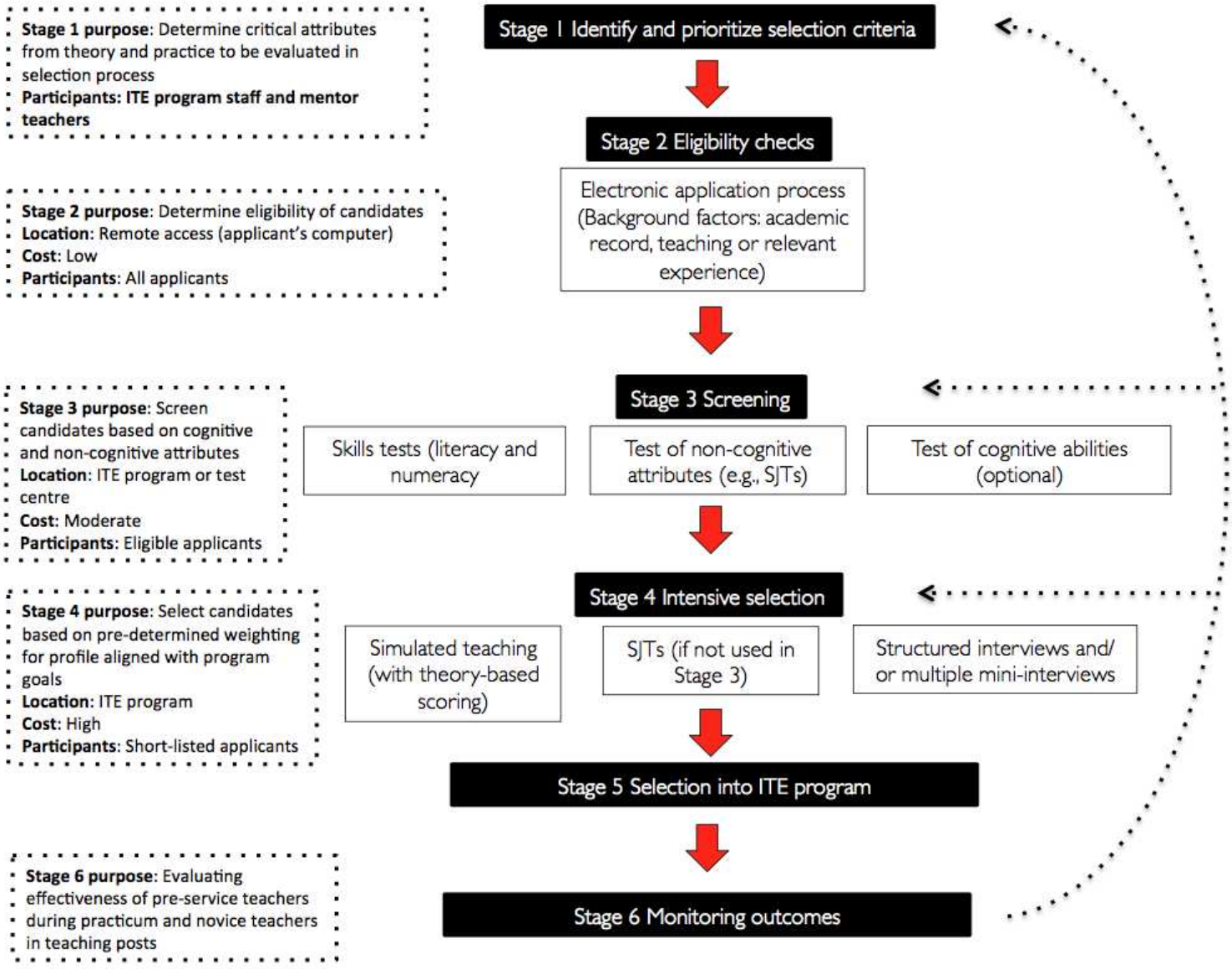


Figure 4. Example of implementation of a six-stage selection process.